FIR System Identification using a Linear Combination of Cumulants

José A. R. Fonollosa, Josep Vidal, and Asunción Moreno

Universitat Politècnica de Catalunya
E.T.S E. Telecomunicació, Apdo 30.002, Barcelona 08080, Spain.

ABSTRACT

In this paper we develop a general linear approach to identify the parameters of a moving average (MA) model from the statistics of the output. First, we show that, under some constraints, the impulse response of the system can be expressed as a linear combination of cumulant slices. Then, this result is used to obtain a new well-conditioned linear method to estimate the MA parameters of a non-gaussian process. The proposed approach does not require a previous estimation of the filter order. Simulation results show improvement in performance with respect to existing methods.

1. INTRODUCTION

Spectral analysis based on higher-order statistics has received great attention in recent years. The developed tools allow to deal with problems where either nonlinearities, non-gaussianity, or nonminimum phase systems are present, and they are of great value in diverse fields as radar, sonar, array processing, blind equalization, time-delay estimation, image and speech processing, and seismology [1,2].

This paper presents a new approach to identify a (possibly) nonminimum phase linear system driven by i.i.d. non-gaussian noise from just output measurements. This problem can only be solved if higher-order statistics are used since, as it is well known, second-order statistics are phase-blind, i. e., they contain only amplitude information. The developed method can also be applied when the measurements are contaminated with additive colored gaussian noise or i.i.d. noise.

Although only systems with a finite impulse response (FIR) are considered, the method we propose can be applied to the identification of the MA parameters of an ARMA model using the AR compensated process (or the AR compensated cumulants). It can also be used to identify both the AR and MA parts of a non-causal ARMA model using the "Double MA Algorithm" [2].

The developed algorithm is linear but it presents several interesting differences with respect to existing linear algebra solutions [2]. These new characteristics can be summarized in three points:

1) The impulse response is computed as a linear combination of cumulants slices. The weights are calculated from an underdetermined linear system of equations and, since any solution to this underdetermined problem gives a consistent estimation, ill-conditioning can be avoided.

2) Different sets of cumulant slices of different orders can be considered. Although it is not necessary, all the cumulants of order 4, order 3 and order 2 (autocorrelation) can be combined. The methods developed by Giannakis and Mendel (GM) [6] and Tugnait (T) [10] require the use of the autocorrelation function while in our algorithm this is just an option,. Consequently, it can provide consistent estimates in the presence of colored gaussian noise of unknown power spectral density. No linear approaches published so far provide a general framework to combine all the statistics.

3) Existing linear methods for estimating the MA parameters as the $C(q,k)$ [5], GM and T methods require an exact knowledge of the MA order and give totally erroneous estimates if the order is overestimated. The proposed algorithm does not need a previous estimation of the filter order or any other parameters as those required by the Bicepstral method [7].

The following sections are a reduced version of paper [8]

2. CUMULANT SLICES

In [4] it was shown that for an AR model, there is always a linear combination of 1-D cumulant slices (w-slice) that gives the impulse response of the system. This w-slice may then be used to compute the AR parameters. In this paper this result is extended to MA processes and a general procedure for computing the weights of the linear combination is given.

Consider a MA(q) process $x(n)$

$$x(n) = \sum_{k=0}^{q} b(k)u(n\text{-}k) \qquad (b(0) := 1,\ b(q) \neq 0) \qquad (1)$$

where $u(n)$ is a driving i.i.d. non-gaussian sequence with $E\{u(n)\}=0$ and m-th order cumulants γ_m. The Brillinger-Rosenblatt summation formula relates the m-th order cumulant of x(n) to the impulse response b(n) as follows

$$C_m(i_1, i_2, ..., i_{m\text{-}1}) = \gamma_m \sum_{n=0}^{q} \prod_{k=0}^{m\text{-}1} b(n+i_k) \qquad i_0 = 0 \qquad (2)$$

Our goal, the FIR system identification problem, is to recover the right side terms b(n) from the left size terms $C_m()$. If we fix the indexes i_2 to $i_{m\text{-}1}$ the resulting 1-D cumulant slice can be expressed as the cross-correlation of $b(n)$ and $b(n; i_2, ..., i_{m\text{-}1})$

$$C_m(i, i_2, ..., i_{m\text{-}1}) = \sum_{n=0}^{q} b(n+i)\, b(n; i_2, ..., i_{m\text{-}1}) \qquad (3)$$

where the causal sequence $b(n; i_2, ..., i_{m\text{-}1})$ is defined as

$$b(n; i_2, ..., i_{m\text{-}1}) := \gamma_m\, b(n) \prod_{k=2}^{m\text{-}1} b(n+i_k) \qquad (4)$$

From (3) and (4) its clear that any general linear combination of slices

$$C_w(i) = w_2\, C_2(i) + \sum_j w_3(j)\, C_3(i, j) +$$
$$+ \sum_{j,k} w_4(j, k)\, C_4(i, j, k) + ... \qquad (5)$$

can also be expressed as the cross-correlation of $b(n)$ and $g_w(n)$

$$C_w(i) = \sum_{n=0}^{q} b(n+i)\, g_w(n) \qquad (6)$$

where $g_w(n)$ is the following causal sequence

$$g_w(n) = w_2\, b(n) + \sum_j w_3(j)\, b(n; j) + ... \qquad (7)$$

Equation (6) shows that, for a MA model, any w-slice can be expressed as the cross-correlation of two finite causal sequences $b(n)$ and $g_w(n)$. The idea behind the developed FIR system identification method is to choose the weights that give $g_w(n) = \delta(n)$ since then $C_w(i)$ will be equal to $b(i)$. Of course, we cannot use (7) to choose the weights because we do not know the sequences $b(i;...)$, but we can use the following results and properties of the w-slices.

Theorem 1: If a w-slice $C_w(i$ is causal, then $C_w(i)=C_w(0)b(i)$.

*Proof:*This theorem and other important properties of the cumulant slices can be proven using basic linear algebra results if we interpret the 1-D slices as vectors:

$$C_m(i_2,...) = (C_m(\text{-}q,i_2,...),...,C_m(0,i_2,...),...,C_m(q,i_2,...))^t$$

$$C_w = (C_w(\text{-}q), ..., C_w(0), ..., C_w(q))^t$$

Then, equation (5) can be expressed in matrix notation as

$$C_w = S\, w \qquad (8)$$

where the matrix of cumulants S and the vector w are defined as

$$S = \begin{pmatrix} C_2(\text{-}q) & \cdots & C_3(\text{-}q, j) & \cdots & C_4(\text{-}q, j, k) & \cdots \\ \vdots & & \vdots & & \vdots & \\ C_2(0) & \cdots & C_3(0, j) & \cdots & C_4(0, j, k) & \cdots \\ \vdots & & \vdots & & \vdots & \\ C_2(q) & \cdots & C_3(q, j) & \cdots & C_4(q, j, k) & \cdots \end{pmatrix}$$

$$w = (w_2...w_3(j)...w_4(j, k) ...)^t$$

Equation (7) can also be rewritten as

$$C_w = \sum_{n=0}^{q} g_w(n)\, b_n \qquad (9)$$

where b_n is a n-shifted version of $b_0=(0,...,0,1,...,b(q))^t$

This equation clearly shows that any slice or w-slice can be expressed as the linear combination of the $q+1$ coefficient vectors b_n. Observe that these vectors are linearly independent. From (9) it is now straightforward to see that if a w-slice is causal, then $g_w(q)=g_w(q-1)=...=g_w(1)=0$, that is, $C_w(i)$ is proportional to b_0.

If $\gamma_m \neq 0$ the slice $C_m(q, 0, ..., 0)$ is a simple example of a causal w-slice since

$$C_m(i, q, ..., 0) = \begin{cases} 0 & i < 0 \\ \gamma_m b(q) \neq 0 & i = 0 \end{cases}$$

It is also a proof of the existence of causal w-slices.

In particular, if a w-slice $C_w(i)$ is causal and $C_w(0)=1$, then $C_w(i)=b(i)$.

The matrix notation used in (8) and (9) allows to obtain other useful properties of the *slices subspace* as the following theorem. The proof can be found in [8].

Theorem 2: The rank of S, i.e., the dimension of the slices subspace, is equal to r, the number of non-zero coefficients.

So far we have assumed that the matrix S was formed by all the cumulant slices of any order. This is the most general assumption since any linear combination of slices is considered. Hence, it is clear that theorem 1 is still valid if only a finite number of slices (rows) is considered. The rank of the corresponding finite matrix S will be also r if a basis is included [8].

3. ALGORITHM

In section II several properties of the w-slices and slices matrix or subspace have been presented. In this section we will study the application of these theoretical results in the development of FIR System Identification methods based on cumulants.

Consider the $MA(q)$ process given by (1). *Corollary 1* indicates that the coefficients $b(i)$ of the FIR filter can be recovered from the cumulants if we find a set of weights that gives a causal w-slice with $C_w(0)=1$. The resulting system of equations can be condensed in a single matrix equation

$$S\,w = b_0 \qquad (11)$$

where the unknowns are the vector w (weights), and the last q elements of b_0 (coefficients). In general the number of unknowns is greater than the number of equations but, provided that the true statistics are used to construct S, *Corollary 1* assures that the solution is unique for the coefficients.

In a practical system identification problem the cumulants have to be estimated. In this case the rank of S will not be *exactly* r and the solution of (11) for both w and the coefficients will not be unique. If the covariances of the cumulant estimates were available, we could think in developing an optimum criterion to choose one solution for the coefficients.

As a first approach we could think in using only one nonzero weight. In fact, in section II it was shown that the slices $C_m(q, 0, ..., 0)$ were causal, so we can use only one of these slices in the linear combination to obtain

$$
\begin{aligned}
b_0 \;&= S\,w = w_m(q, 0, ..., 0)\,C_m(q, 0, ..., 0) = \\
&\frac{1}{C_m(0, q, 0, ..., 0)}\,C_m(q, 0, ..., 0)
\end{aligned} \qquad (12)
$$

The resulting equation is nothing else but the $C(q,k)$ method [5]. This is an important theoretical result. Nevertheless, in practice, the $C(q,k)$ method does not provide good estimates since only one slice is considered. Instabilities are also likely to appear due to the estimated cumulant $C_m(0, q, 0, ..., 0)$ may be close to *zero*.

The proposed approach to combine all the slices in the estimation is to choose the vector w with the minimum norm. This is a simple and usual solution to a underdetermined system equations and, if SVD is used, the solution will be always well conditioned. The results

of the simulations confirm the good behavior of this method. In fact, this solution would be the optimum solution if we did not consider the error in the estimation of the weights and the cumulants estimates were independent and had all the same variance. The resulting algorithm is the following:

W-Slice Algorithm. (WS) The matrix equation (11) is solved in two steps.

S1) Computation of the minimum-norm weights that give a causal w-slice with $C_w(0)=1$.

$$S_u\,w = (0, ..., 0, 1)^t = 1$$

$$w_m = S_u^{\#}\,1 \qquad (13)$$

where S_u is the matrix formed with the upper $q+1$ rows of S and $S_u^{\#}$ denotes its pseudoinverse.

S2) Computation of the coefficients as $b(i)=C_w(i)$, i.e.,

$$b_0 = S\,w_m = S\,S_u^{\#}\,1 \qquad (14)$$

Observe that *theorem 1* is still valid if b(q) is zero. Hence, if SVD is used to obtain the minimum-norm weights we do not require an exact knowledge of the order q or the number of nonzero coefficients r. The $C(q,k)$ method is also a proof that we do not require a complete basis to obtain a consistent estimator. If the order q is known we just need to include the slice $C_m(q, 0, ..., 0)$ in S to assure the consistency of the WS estimation. If we only have an upper bound (q_{max}) and a lower bound (q_{min}) of q then the required minimum set of slices for the WS algorithm is

$$C_m(j, 0, ..., 0) \qquad q_{min} \leq j \leq q_{max}$$

Observe it is not necessary to include the $C_m(0, 0, ..., 0)$ slice to obtain a consistent estimator. Hence, we can still use this approach in the presence of non-gaussian i.i.d. noise. Of course, any kind of gaussian noise is not a problem either, since the slice C_2 is not necessary in any case.

4. EXAMPLES

The objective of the simulations were to compare the performance of the proposed algorithm with existing approaches and to check the good behavior of the algorithm when the system order was overestimated.

We present here the results obtained with the WS algorithm and the *Modification to Reformulated GM Algorithm* proposed by Tugnait in [11]. In the WS algorithm the $2Q+1$ third-order cumulant slices $C_3(j)$ $(j = -Q, ..., Q)$ have been used to form a $2Q+1$ by $2Q+1$ statistics matrix S.

Example 1: In this example we consider the MA(1) and MA(2) processes studied in [3]. The input is an i.i.d. (one-

sided) exponentially distributed random sequence. No noise is added to the output samples used to estimate the parameters. For the WS algorithm Q is equal to the true order q. The results of a Monte Carlo simulation with 1000 runs are shown in table I (mean ± standard deviation). The record length was 400 for the MA(1) models and 1000 for the MA(2) models.

Table I. Example 1, 1000 Monte Carlo runs, N=400 / 1000.

Parameter	True value	GMT	WS
b(1)	-0.80	-0.79 ± 0.06	-0.80 ± 0.03
b(1)	-1.25	-1.23 ± 0.11	-1.23 ± 0.11
b(1)	-1.40	-1.41 ± 0.29	-1.41 ± 0.16
b(2)	0.98	0.97 ± 0.11	0.99 ± 0.17
b(1)	-1.13	-1.12 ± 0.16	-1.13 ± 0.09
b(2)	0.60	0.59 ± 0.11	0.61 ± 0.07

Example 2: The MA(1) processes of example 1 are considered in this example to test the WS algorithm when the order is overestimated. Since none of the other algorithms works in this case, only the results for the WS algorithm are presented in Table II. The Monte Carlo runs were performed in the same conditions. Although there is an increase in the variance of the estimation, both *b(1)* and *b(2)* are still correctly estimated.

Table II. Example 3, 1000 Monte Carlo runs, N=400.

Parameter	True value	WS (Q=2)
b(1)	-0.80	-0.79 ± 0.08
b(2)	0.00	0.00 ± 0.09
b(1)	-1.25	-0.22 ± 0.11
b(2)	0.00	-0.01 ± 0.12

Example 3: In this case we have considered the same MA(5) process used in [10, 11], where several methods of MA parameter estimation were compared. Our results for the C(q,k), GMT1, and GMT2 algorithms are in agreement with those presented in [10,11] for a record length of 5120 samples. Table III shows that the proposed WS method outperforms all the other linear methods in both bias and standard deviation.

Table IV. Example 4, 128 Monte Carlo runs, N=5120, q=5.

Parameter	True value	GMT	WS
b(1)	0.100	0.01 ± 0.54	0.14 ± 0.31
b(2)	-1.870	-1.33 ± 0.51	-1.98 ± 0.60
b(3)	3.020	2.05 ± 0.70	3.05 ± 0.67
b(4)	-1.435	-0.72 ± 0.60	-1.44 ± 0.33
b(5)	0.490	0.22 ± 0.29	0.47 ± 0.17

5. CONCLUSIONS

This paper studied the properties of the cumulant slices of a MA process. This properties of the slices subspace were used to develop a general approach for the identification of FIR systems for output measurements. The proposed algorithm can be based on different sets of cumulant slices and it allows to obtain consistent estimates when colored gaussian noise and/or i.i.d. noise is present in the measurements. The simulations also showed that it gives competitive results in bias and variance, and that it has a good behavior even if the order is overestimated.

REFERENCES

[1] C.L. Nikias and M.R. Raghuveer, "Bispectrum estimation: a digital signal processing framework", *Proc. IEEE*, vol. 75, pp. 869-891, July 1987.

[2] J. M. Mendel, "Tutorial on higher-order statistics (spectra) in signal processing and system sheory: theoretical results and some applications", *Proc. IEEE*, vol. 79, no. 3, pp. 278-305, March 1991.

[3] B. Porat and B. Friedlander, "Performance analysis of parameter estimation algorithms based on high-order moments". *Int. J. Adaptive Control and Signal Processing*, vol. 3, pp. 191-229. 1989.

[4] B. Friedlander and B. Porat, "Adaptive IIR algorithms based on high-order statistics". *IEEE Trans. Acoust. Speech, Signal Processing*, vol. 37, pp. 485-495.

[5] G. B. Giannakis, "Cumulants: a powerful tool in signal processing". *Proc. IEEE*, vol. 75, pp. 1333-1334, 1987.

[6] G. B. Giannakis and J. M. Mendel, "Identification of non-minimum phase systems using higher-order statistics", *IEEE Trans. Acoust. Speech, Signal Processing*, vol. 37, pp. 360-377. 1989.

[7] R. Pan. and C. L. Nikias, "The complex cepstrum of higher-order moments", *IEEE Trans. Acoust. Speech, Signal Processing*, vol. 36, pp. 186-205, 1988.

[8] J. A. R. Fonollosa, E. Vidal, "System Identification using a linear combination of slices", submitted for publication.

[9] J. A. R. Fonollosa, E. Vidal, and E. Masgrau, "Adaptive system identification based on higher-order statistics", in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Processing*, pp. 3437-3440, May 1991.

[10] J. K. Tugnait, "Approaches to FIR system identification with noisy data using higher order statistics", *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 38, no. 7, pp 1307-1317, July 1990.

[11] J. K. Tugnait, "New results on FIR system identification using higher order statistics", *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 39, no. 10, pp 2216-2221, October 1991.